০০১-০০১ ডেটা সায়েন্সের ভূমিকা

ডেটা কি?

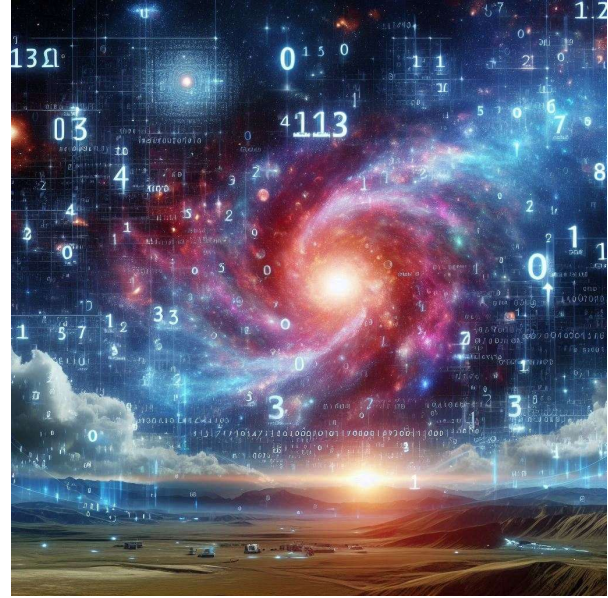
ডেটা (Data) হল কোনো নির্দিষ্ট প্রসঙ্গে সংগ্রহ করা কাঁচা তথ্য, সংখ্যা, অক্ষর, চিত্র বা যেকোনো উপাত্ত যা বিশ্লেষণ করা সম্ভব। এটি মূলত অব্যবহৃত বা অপরিবর্তিত তথ্য, যা বিভিন্ন সূত্র থেকে সংগৃহীত হয়।

উদাহরণ:

- ছাত্রদের পরীক্ষার নম্বর
- কোনো ওয়েবসাইটে ভিজিটরদের সংখ্যা
- তাপমাত্রার হিসাব

ডেটা সায়েন্স কি?

ডেটা সায়েন্স (Data Science) হল ডেটা বিশ্লেষণ, প্রক্রিয়াকরণ ও ব্যাখ্যা করার একটি আধুনিক বিজ্ঞান যা বিভিন্ন অ্যালগরিদম, পরিসংখ্যান এবং মেশিন লার্নিং ব্যবহার করে কার্যকর সিদ্ধান্ত নিতে সহায়তা করে।

এর মূল লক্ষ্য হলো বিশাল পরিমাণের ডেটা থেকে অর্থবহ তথ্য বের করে ব্যবসা, গবেষণা, এবং প্রযুক্তি ক্ষেত্রে উন্নতি করা।

ইনফরমেশন কি?

ইনফরমেশন (Information) হল প্রক্রিয়াজাত ও বিশ্লেষণ করা ডেটা যা মানুষের জন্য অর্থবহ এবং সিদ্ধান্ত গ্রহণে সহায়ক। ডেটার ওপর বিভিন্ন বিশ্লেষণ ও প্রক্রিয়াকরণের মাধ্যমে যখন তা অর্থবহ হয়, তখন তাকে ইনফরমেশন বলা হয়।

উদাহরণ: - একজন ছাত্রের পরীক্ষার নম্বর (ডেটা), কিন্তু তার সর্বোচ্চ নম্বর, গড় নম্বর বা পাস/ফেল হওয়ার তথ্য (ইনফরমেশন)।

ডেটা সায়েন্সের ব্যবহারসমুহ:

ডেটা সায়েন্সের ব্যবহার বিভিন্ন ক্ষেত্রে দেখা যায়, যেমন:

1. ব্যবসা ও মার্কেটিং: গ্রাহকদের কেনার অভ্যাস বিশ্লেষণ করে বিজ্ঞাপন তৈরি করা।

2. স্বাস্থ্য সেবা: রোগ নির্ণয় এবং চিকিৎসা পূর্বাভাস তৈরি করা।

3. আর্থিক খাত: জালিয়াতি শনাক্তকরণ এবং শেয়ার বাজার পূর্বাভাস।

4. সোশ্যাল মিডিয়া: ইউজার বিহেভিয়ার বিশ্লেষণ করে কাস্টমাইজড কন্টেন্ট দেখানো।

5. স্বয়ংক্রিয় যানবাহন: স্বয়ংক্রিয় গাড়ির (self-driving cars) সিদ্ধান্ত গ্রহণ।

6. স্পোর্টস অ্যানালিটিক্স: খেলোয়াড়দের পারফরম্যান্স বিশ্লেষণ ও কৌশল নির্ধারণ।

<u>ডেটা সায়েন্সের প্রয়োজনীয়তা:</u>

বর্তমান বিশ্বে ডেটা সায়েন্স অত্যন্ত গুরুত্বপূর্ণ কারণ এটি বিভিন্ন সেক্টরে সিদ্ধান্ত গ্রহণকে সহজ ও কার্যকরী করে তোলে।

এর প্রয়োজনীয়তা নিম্নলিখিত কারণে বাড়ছে:

- বিশাল পরিমাণের ডেটা সংগ্রহ ও ব্যবহারের প্রয়োজন।

- ব্যবসা ও প্রযুক্তির অগ্রগতি।

- স্বাস্থ্য, শিক্ষা ও গবেষণার উন্নতির জন্য বিশ্লেষণমূলক সিদ্ধান্ত গ্রহণ।

- আর্টিফিশিয়াল ইন্টেলিজেন্স (AI) এবং মেশিন লার্নিং (ML) নির্ভর প্রযুক্তির বৃদ্ধি।

<u>ডেটা সায়েন্সের ডিমান্ড:</u>

বর্তমানে ডেটা সায়েন্স পেশার চাহিদা দ্রুতগতিতে বাড়ছে, কারণ:

1. বড় কোম্পানিগুলো (Google, Facebook, Amazon, Microsoft) বিশাল পরিমাণ ডেটা ব্যবহার করে।

2. বিভিন্ন দেশ ও প্রতিষ্ঠান ডেটা-ড্রিভেন সিদ্ধান্ত গ্রহণ করছে।

3. ডেটা সায়েন্টিস্টদের উচ্চ বেতনের চাকরি পাওয়ার সম্ভাবনা বেশি।

4. মেশিন লার্নিং ও AI-এর বিকাশের সাথে সাথে ডেটা সায়েন্স আরও গুরুত্বপূর্ণ হয়ে উঠছে

বর্তমানে ডেটা সায়েন্স দক্ষতা সম্পন্ন পেশাজীবীদের প্রচুর চাহিদা রয়েছে এবং এটি ভবিষ্যতেও বৃদ্ধি পাবে।

<u>ডেটা সায়েন্স শেখার জন্য যা যা দরকারঃ</u>

১) Statistics শেখা লাগবে।

২) পাইথন প্রোগ্রামিং জানা লাগবে।

এই দুইটার কম্বিনেশন হল ডেটা সায়েন্স।

বিঃদ্রঃ অনেকের মনে হতে পারে যে পাইথন বাদে তো R বা STATA দিয়ে ডেটা সায়েন্সের প্রোগ্রাম করা যায়। তবে সমস্যা হলো যে STATA হলো paid (expensive license) এর প্রোগ্রাম, সেখানে ডেটা ভিজুয়ালাইজেশন লিমিটেড; মেশিন লার্নিং ভেরি লিমিটেড; বিগ ডেটা, ডীপ লার্নিং এবং এআই এর জন্য কার্যকর নয়। R মেশিন লার্নিং এর ক্ষেত্রে লিমিটেড; বিগ ডেটার ক্ষেত্রে মডারেট; ডীপ লার্নিং এবং এআই এর ক্ষেত্রে দূর্বল। পাইথনে এসকল কোনো সমস্যা নেই। তবে Statistical analysis আর ভিজুয়ালাইজেশনের ক্ষেত্রে R অনেকটা ভালো। এখানে আমরা পাইথন দিয়ে ডেটা সায়েন্সের প্রোগ্রামিং শিখবো। আর যদি সুযোগ হয়, R প্রোগ্রামিং দেখতে পারব, ইন শা আল্লাহ।